nature genetics

Mutation of *SHOC2* promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair

Viviana Cordeddu[1], Elia Di Schiavi[2], Len A Pennacchio[3,4], Avi Ma'ayan[5], Anna Sarkozy[6], Valentina Fodale[1,7], Serena Cecchetti[8], Alessio Cardinale[9], Joel Martin[4], Wendy Schackwitz[4], Anna Lipzen[4], Giuseppe Zampino[10], Laura Mazzanti[11], Maria C Digilio[12], Simone Martinelli[1], Elisabetta Flex[1], Francesca Lepri[6], Deborah Bartholdi[13], Kerstin Kutsche[14], Giovanni B Ferrero[15], Cecilia Anichini[16], Angelo Selicorni[17], Cesare Rossi[18], Romano Tenconi[19], Martin Zenker[20], Daniela Merlo[8,9], Bruno Dallapiccola[6,7], Ravi Iyengar[5], Paolo Bazzicalupo[2], Bruce D Gelb[21,22] & Marco Tartaglia[1,22]

N-myristoylation is a common form of co-translational protein fatty acylation resulting from the attachment of myristate to a required N-terminal glycine residue[1,2]. We show that aberrantly acquired N-myristoylation of SHOC2, a leucine-rich repeat–containing protein that positively modulates RAS-MAPK signal flow[3–6], underlies a clinically distinctive condition of the neuro-cardio-facial-cutaneous disorders family. Twenty-five subjects with a relatively consistent phenotype previously termed Noonan-like syndrome with loose anagen hair (MIM607721)[7] shared the 4A>G missense change in *SHOC2* (producing an S2G amino acid substitution) that introduces an N-myristoylation site, resulting in aberrant targeting of SHOC2 to the plasma membrane and impaired translocation to the nucleus upon growth factor stimulation. Expression of SHOC2^{S2G} *in vitro* enhanced MAPK activation in a cell type–specific fashion. Induction of SHOC2^{S2G} in *Caenorhabditis elegans* engendered protruding vulva, a neomorphic phenotype previously associated with aberrant signaling. These results document the first example of an acquired N-terminal lipid modification of a protein causing human disease.

Dysregulation of the RAS-MAPK signaling pathway has recently been recognized as the molecular cause underlying a group of clinically related developmental disorders with features including reduced growth, facial dysmorphism, cardiac defects, ectodermal anomalies, variable cognitive deficits and susceptibility to certain malignancies[8,9]. These mendelian traits are caused by mutations in genes encoding RAS proteins (*KRAS* and *HRAS*), downstream transducers (*RAF1*, *BRAF*, *MEK1* and *MEK2*) or pathway regulators (*PTPN11*, *SOS1*, *NF1* and *SPRED1*). For Noonan syndrome, the commonest of these disorders, mutations are observed in several of these genes, underlying approximately 70% of cases.

To rationalize further candidate gene approaches to Noonan syndrome gene discovery, we used a systems biology approach based on *in silico* protein network analysis. By applying a graph theory algorithm on a filtered consolidated human interactome, we derived a subnetwork of proteins generated from an integrated network of mammalian protein interaction databases and cell-signaling network datasets by seeding with the known disease-causing mutant proteins (**Supplementary Fig. 1a**). To identify potential Noonan syndrome disease genes, we computed *Z* scores using a binomial proportions test, which ranked the significance of the intermediate nodes within the subnetwork based on their connections to the seed proteins[10] (**Supplementary Table 1**). We resequenced coding exons for the best gene candidate, *SHOC2*, in a Noonan syndrome cohort that included 96 individuals who were negative for mutations in known

[1]Dipartimento di Ematologia, Oncologia e Medicina Molecolare, Istituto Superiore di Sanità, Rome, Italy. [2]Istituto di Genetica e Biofisica 'A. Buzzati Traverso', Consiglio Nazionale delle Ricerche, Naples, Italy. [3]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [4]US Department of Energy Joint Genome Institute, Walnut Creek, California, USA. [5]Department of Pharmacology and Systems Therapeutics, Systems Biology Center New York (SBCNY), Mount Sinai School of Medicine, New York, New York, USA. [6]Istituto di Ricovero e Cura a Carattere Scientifico-Casa Sollievo della Sofferenza, San Giovanni Rotondo and Istituto Mendel, Rome, Italy. [7]Department of Experimental Medicine, University 'La Sapienza', Rome, Italy. [8]Dipartimento di Biologia Cellulare e Neuroscienze, Istituto Superiore di Sanità, Rome, Italy. [9]IRCCS-San Raffaele Pisana, Rome, Italy. [10]Istituto di Clinica Pediatrica, Università Cattolica del Sacro Cuore, Rome, Italy. [11]Dipartimento di Pediatria, Università degli Studi di Bologna, Bologna, Italy. [12]Sezione di Genetica Medica, Ospedale Bambino Gesù, Rome, Italy. [13]Institute of Medical Genetics, University of Zurich, Schwerzenbach, Switzerland. [14]Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany. [15]Dipartimento di Pediatria, Università di Torino, Turin, Italy. [16]Dipartimento di Pediatria, Ostetricia e Medicina della Riproduzione, Università di Siena, Siena, Italy. [17]Clinica Pediatrica I, IRCCS Fondazione Policlinico Milano, Milan, Italy. [18]Unità Operativa di Genetica Medica, Policlinico S. Orsola-Malpighi, Bologna, Italy. [19]Dipartimento di Pediatria, Università di Padova, Padua, Italy. [20]Institute of Human Genetics, University Hospital Erlangen, University of Erlangen-Nuremberg, Erlangen, Germany. [21]Center for Molecular Cardiology and Departments of Pediatrics and Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA. [22]These authors contributed equally to this work. Correspondence should be addressed to M.T. (mtartaglia@iss.it) or B.D.G. (bruce.gelb@mssm.edu).

Received 1 April; accepted 22 June; published online 16 August 2009; doi:10.1038/ng.425

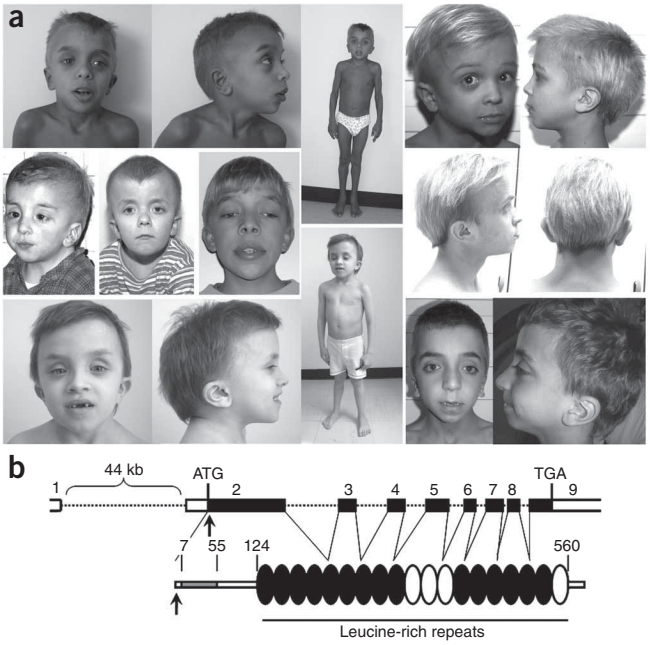

Figure 1 The germline 4A>G mutation in the *SHOC2* gene causes a distinctive phenotype of the neuro-cardio-facial-cutaneous syndrome family. (**a**) Representative phenotypic features of affected subjects carrying the *SHOC2* mutation. Common features include macrocephaly, high forehead, hypertelorism, palpebral ptosis, low-set and posteriorly rotated ears, short and webbed neck, and pectus anomalies. Affected subjects also had easily pluckable, sparse, thin, slow-growing hair. (**b**) *SHOC2* genomic organization and protein structure. The coding exons are shown at the top as numbered filled boxes. Intronic regions are represented by dotted lines. SHOC2's motifs comprise an N-terminal lysine-rich region (in gray; Prosite motif score = 8.8) followed by 19 leucine-rich repeats (Pfam hits with an *E* value <0.5 are in black and those with an *E* value >1 in white). Numbers above the domain structure indicate the amino acid boundaries of those domains.

disease genes and who were opportunely selected to represent the wide phenotypic spectrum characterizing Noonan syndrome; this revealed an A-to-G transition at position 4 of the gene, predicting the substitution of a glycine for serine at position 2 (S2G), in four unrelated individuals (**Supplementary Fig. 1b**). All cases were diagnosed as sporadic, and genotyping of parental DNA available for three of the four subjects documented the absence of the sequence variant and confirmed paternity in each family, providing evidence that the change was a *de novo* mutation associated with the disease. For these subjects, DNA samples from several tissues were available, and all harbored the S2G mutation, providing evidence that the defect was inherited through the germline. We then analyzed *SHOC2* in a cohort of 410 mutation-negative subjects with Noonan syndrome or a related phenotype. We observed 21 individuals with the 4A>G missense change and proved that the causative mutations were *de novo* in 12 families from which parental DNA was available. No additional disease-associated *SHOC2* sequence variant was identified in this cohort, strongly suggesting a specific pathogenetic role for the S2G amino acid substitution.

Review of the clinical features of individuals with the *SHOC2* mutation revealed a consistent phenotype, previously termed Noonan-like syndrome with loose anagen hair[7]. Although their facial features seemed typical of Noonan syndrome (**Fig. 1a**), phenotypic analysis of these subjects was notable in that they showed reduced growth that was frequently associated with proven growth hormone deficiency, cognitive deficits, distinctive hyperactive behavior that improved with age in most subjects, and hair anomalies including easily pluckable, sparse, thin, slow-growing hair. In 12 subjects, a diagnosis of loose anagen hair was confirmed by microscopic examination of pulled hairs. Most of the subjects also had darkly pigmented skin with eczema or ichthyosis. Cardiac anomalies were observed in the majority of the subjects, with dysplasia of the mitral valve and septal defects considerably overrepresented compared with the general population of individuals with Noonan syndrome. The affected individuals' voices were characteristically hypernasal. Of note, the referring pediatricians felt that several of these subjects had features suggestive of Costello syndrome or of cardiofaciocutaneous syndrome as newborns or young infants. Overall, these subjects appeared to share a phenotype that was characterized by an unusual combination of features observed in disorders of the neuro-cardio-facial-cutaneous disorders family (**Supplementary Table 2**).

SHOC2 is a widely expressed protein composed almost entirely of leucine-rich repeats (LRR), with a lysine-rich sequence at the N terminus (**Fig. 1b**). In *C. elegans*, where SHOC2 (also called SUR-8 and SOC-2) was discovered, the SHOC2 protein acts as a positive modulator of the RAS-MAPK signaling cascade, which is elicited by EGL-15 and LET-23 and mediated by LET-60 (homologs of vertebrate FGFR, EGFR and RAS family members, respectively[3,4]). Because LRRs can provide a structural framework for protein-protein interactions, SHOC2 is believed to function as a scaffold, linking RAS to downstream signal transducers[4–6]. Based on the N-terminal position of the S2G substitution, we hypothesized that co-translational processing might be perturbed in the mutant protein, making it a substrate for N-myristoyltransferase (NMT). N-terminal myristoylation is an irreversible modification generally occurring during protein synthesis, in which myristate, a 14-carbon saturated fatty acid, is covalently added to an N-terminal glycine residue after excision of the initiator methionine residue by methionyl aminopeptidase[1,2]. For this to occur, glycine at codon 2 is absolutely required, small uncharged residues at positions 3 and 6 are generally needed, and basic residues at positions 7–9 are preferred[11]. Except for the presence of serine at position 2, the N-terminal sequence of SHOC2 satisfies those consensus rules, and *in silico* analysis predicted myristoylation of the $SHOC2^{S2G}$ mutant with high confidence. To verify this, we evaluated the myristoylation status of wild-type and mutant SHOC2 proteins transiently expressed in Cos-1 cells (**Fig. 2a**). We found that $SHOC2^{S2G}$ incorporated [^{3}H]myristic acid, whereas the wild-type protein and the disease-unrelated $SHOC2^{S2A}$ did not.

N-myristoylation facilitates the anchoring of proteins to cellular membranes. To explore whether this process conferred membrane targeting to mutant SHOC2, the subcellular localization of V5-tagged SHOC2 proteins was analyzed in Cos-1 cells (**Fig. 2**). Confocal laser scanning microscopy analysis showed that $SHOC2^{wt}$ was uniformly distributed in the cytoplasm and nucleus in starved cells, but that it was restricted to the nucleus following epidermal growth factor (EGF) stimulation, implying an unexpected role for this protein in signal transduction. In contrast, $SHOC2^{S2G}$ was specifically targeted to the cell membrane both in starved cells and after EGF stimulation. This aberrant localization of $SHOC2^{S2G}$ was confirmed in 293T and Neuro2A cell lines and using a Myc-tagged protein (data not shown) as well as by cell fractionation (**Fig. 2c**). Growth factor–stimulated nuclear translocation of the endogenous SHOC2 protein was confirmed in primary skin fibroblasts (**Fig. 2e**). Treatment with 2-hydroxymyristic acid, an NMT inhibitor, at varying doses reduced or abolished the membrane localization of $SHOC2^{S2G}$ (**Supplementary Fig. 2**), confirming a dependency upon myristoylation. In addition,

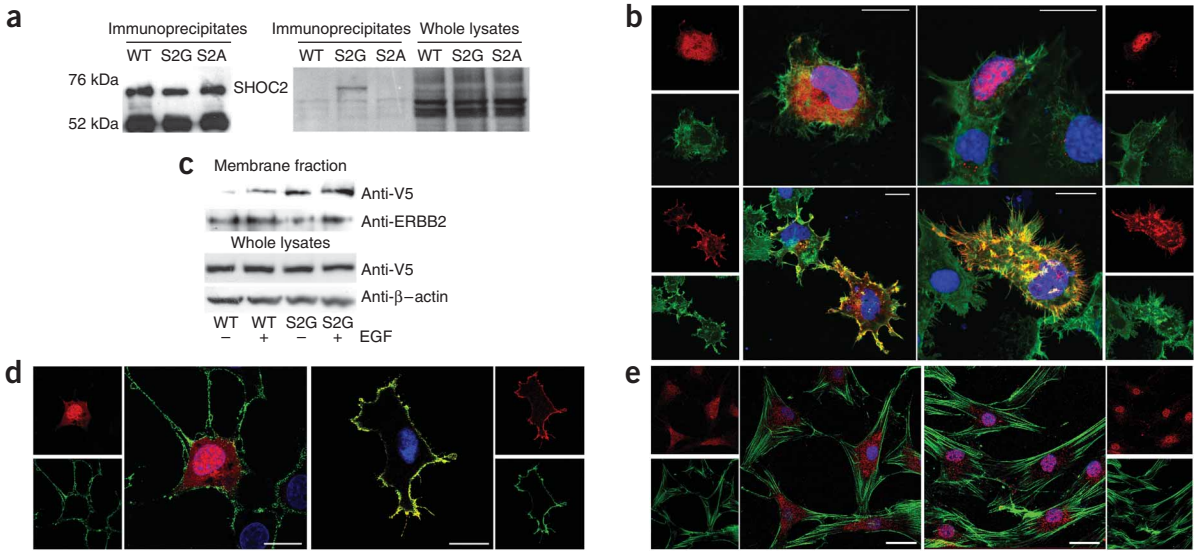

Figure 2 The disease-causing 4A>G change in *SHOC2* promotes protein myristoylation and cell membrane targeting. (**a**) [^{3}H]Myristic acid incorporation (middle) occurs in SHOC2^{S2G} but not in SHOC2^{wt} or SHOC2^{S2A}. Equivalent levels of SHOC2 proteins in immunoprecipitates (left) and [^{3}H]myristic acid incorporation in cells (right) are shown. (**b**) SHOC2^{wt} is uniformly present in the cytoplasm and nucleus in starved Cos-1 cells (upper left) and is restricted to the nucleus following EGF stimulation (upper right), whereas SHOC2^{S2G} is targeted to the cell membrane basally (lower left) and after stimulation (lower right). Confocal microscopy visualized SHOC2 (anti-V5 monoclonal, then Alexa Fluor 594–goat anti-mouse; red), actin cytoskeleton (Alexa Fluor 488–phalloidin; green) and nuclei (DAPI; blue). (**c**) Cell fractioning assay documenting preferential membrane targeting of SHOC2^{S2G}. The protein ERBB2 is shown to demonstrate equivalent fractionation efficiency, whereas anti-V5 blots from cell lysates show equivalent transfection efficiency. (**d**) Colocalization of V5-tagged SHOC2^{S2G} and ganglioside M1 to the plasma membrane in Cos-1 cells. Subcellular localization of V5-tagged wild-type SHOC2 (left) and V5-tagged SHOC2^{S2G} (right) is shown. Ganglioside M1 was detected by using the Vybrant Lipid Raft Labeling kit (green). SHOC2 proteins and nuclei are visualized as reported above. (**e**) Subcellular localization of the endogenous SHOC2^{wt} protein in primary skin fibroblasts, basally (left) and after stimulation (right). Confocal microscopy was performed using an anti-SHOC2 polyclonal followed by Alexa Fluor 594–goat anti-rabbit (red), whereas the actin cytoskeleton was detected by Alexa Fluor 488–phalloidin (green). All images are single optical sections representative of >50 transfected cells observed in each experiment. Bars, 20 μm (**b**,**d**) or 40 μm (**e**).

even in the absence of efficient myristoylation, SHOC2^{S2G} did not translocate to the nucleus upon stimulation, indicating possible loss of function. Efficient nuclear translocation was observed for the disease-unrelated SHOC2^{S2A} mutant following EGF stimulation (**Supplementary Fig. 3**), suggesting a specific effect of the mutation that causes disease. To exclude the possibility that SHOC2^{S2G} might exert a dominant negative effect by sequestering the wild-type protein to the cell membrane, thereby impairing its EGF-dependent translocation to the nucleus, we assayed SHOC2^{wt} and SHOC2^{S2G} heterodimerization by confocal microscopy and coimmunoprecipitation assays in Cos-1 cells transiently co-transfected with V5- and Myc-tagged proteins (**Fig. 3a**,**b**). The experiments demonstrated that these proteins do not heterodimerize, ruling out the possibility of a dominant negative effect by SHOC2^{S2G}. Next, we explored whether SHOC2^{S2G} altered signaling through MAPK by expressing SHOC2^{wt} or SHOC2^{S2G} in Cos-1, 293T and Neuro2A cells. Although we did not observe significant change in ERK activation in Cos-1 and 293T cells (data not shown), SHOC2^{S2G} expression promoted enhanced EGF-dependent ERK phosphorylation compared to wild-type SHOC2 in neuroblastoma Neuro2A cells (**Fig. 3c**,**d**).

To further explore the functional effects of the SHOC2^{S2G} mutant, we used *C. elegans* as an experimental model. In *C. elegans*, reduced SUR-8 function (*sur-8^{rf}*) causes no identifiable phenotype by itself but can suppress the gain-of-function LET-60 (*let-60^{gof}*)-induced multivulva phenotype (Muv)[4]. We tested whether expression of SHOC2 proteins could rescue the suppressed Muv phenotype in the *sur-8^{rf}*, *let-60^{gof}* genetic background (**Supplementary Table 3**). Although SHOC2^{wt} was able to replace SUR-8 functionally, SHOC2^{S2G} failed to do so. Expression of the mutant in *let-60^{gof}* worms did not suppress the Muv phenotype (**Supplementary Table 3**), excluding dominant negative effects for SHOC2^{S2G}. In worms with a wild-type genetic background, SHOC2^{S2G} expression at embryonic and early larval stages of development caused no visible phenotype. In contrast, at the early L3 stage, it caused abnormal vulval development, resulting in protruding vulva (Pvl), decreased egg laying efficiency (Egl) and accumulation of larvae inside the mother with the formation of bag-of-worms adults (Bag phenotype) (**Supplementary Table 4** and **Fig. 4**). These neomorphic phenotypes were absent in animals expressing SHOC2^{wt} but were observed when SHOC2^{wt} tagged with an N-myristoylation sequence (myr::SHOC2^{wt}) was expressed (**Supplementary Table 4** and **Fig. 4**). The SHOC2^{S2G} and myr::SHOC2^{wt} proteins were targeted to the cell membrane in various *C. elegans* cell types, whereas SHOC2^{wt} was observed diffusely throughout the cytoplasm and nucleus (**Fig. 4e–j**). The defects in vulva formation were not due to increased induction of the vulva cell fate in vulval precursor cells (VPC), as expression of SHOC2^{S2G} did not reduce the penetrance of the vulvaless phenotype of a *let-23^{rf}* hypomorph mutant (**Supplementary Table 5**), nor did it increase the penetrance of the Muv phenotype of *let-60^{gf}* animals (**Supplementary Table 3**). At the late L3 to early L4 stage, vulva morphogenesis normally begins with the descendants of VPC P6.p detaching from the cuticle and forming a symmetric invagination. Animals in which the expression of SHOC2^{wt} had been induced at early L3 maintained this pattern. In contrast, in larvae expressing SHOC2^{S2G} (17/48) or myr::SHOC2^{wt} (10/22), descendants of VPCs P5.p and/or P7.p also detached from the cuticle, resulting in larger and more

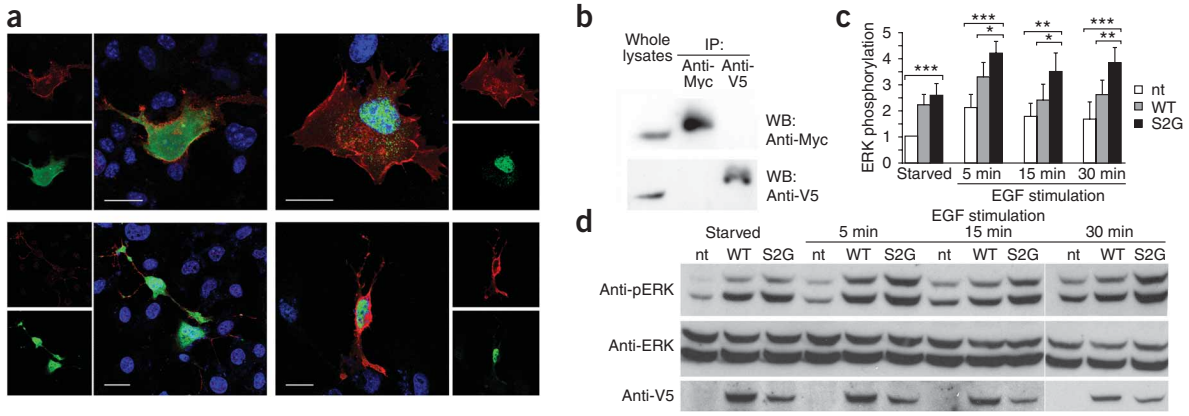

Figure 3 Functional characterization of the disease-causing 4A>G change in *SHOC2*. (**a**) Subcellular localization of coexpressed SHOC2^{wt} (green) and SHOC2^{S2G} (red) documenting that SHOC2^{S2G} does not impair EGF-stimulated SHOC2^{wt} translocation to the nucleus. Imaging of V5-tagged (anti-V5 monoclonal, then Alexa Fluor-594 goat anti-mouse) and Myc-tagged (anti-Myc, then Alex Fluor 488 goat anti-rabbit) SHOC2 proteins and nuclei (DAPI, blue). Panels above show Myc-tagged SHOC2^{wt} and V5-tagged SHOC2^{S2G}, and those below show V5-tagged SHOC2^{wt} and Myc-tagged SHOC2^{S2G}. Cells were imaged basally (left) and after EGF stimulation (right). Bars indicate 20 μm. (**b**) Lysates of Cos-1 cells coexpressing Myc-tagged SHOC2^{wt} and V5-tagged SHOC2^{S2G} were immunoprecipitated using anti-Myc (above) or anti-V5 (below), and immunoprecipitated proteins were visualized by protein blotting. These results indicate that SHOC2 proteins do not form heterodimers. (**c**,**d**) ERK phosphorylation in V5-tagged SHOC2^{wt} or SHOC2^{S2G} transiently expressed Neuro2A cells basally or following EGF stimulation. Phosphorylation levels are reported as a multiple of basal ERK phosphorylation in cells not transfected with a SHOC2 construct, averaged from four replicates ± s.d. (**c**). Results for cells expressing the S2G mutant were compared with those overexpressing the wild-type protein (below) or untransfected cells (above) at the same time points using two-tailed *t*-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Representative blots are also shown (**d**).

asymmetric invaginations (**Fig. 4k–n**). This morphogenesis defect was the earliest detectable neomorphic effect of the SHOC2^{S2G} mutation on vulval development.

We discovered that a *SHOC2* mutation promoting N-myristoylation of its protein product causes Noonan-like syndrome with loose anagen hair. This acquired fatty acid modification is unique in inherited human disease and results in constitutive membrane targeting that leads to increased MAPK activation in a cell context–specific manner. Cell-specific activation of the RAS pathway has also been observed in Noonan syndrome–associated SHP-2 mutants[12–14]. Although not well understood, this phenomenon explains why, despite the ubiquitousness of RAS signaling, development is perturbed in certain tissues in these disorders. It has recently been reported that SHOC2 functions as a regulatory subunit of the catalytic subunit of protein phosphatase 1 (PP1C)[6]. By binding GTP-MRAS, SHOC2 promotes PP1C translocation to the membrane, allowing PP1C-mediated RAF1 dephosphorylation at residue Ser259, which is required for stable RAF1 translocation to the plasma membrane and catalytic activation. Of note, the portion of *RAF1* encoding Ser259 and adjacent residues, which represent a 14-3-3 protein binding site with inhibitory function, is a major hot spot for Noonan syndrome–causing mutations affecting RAF1 (ref. 9). According to this model, constitutive membrane translocation of

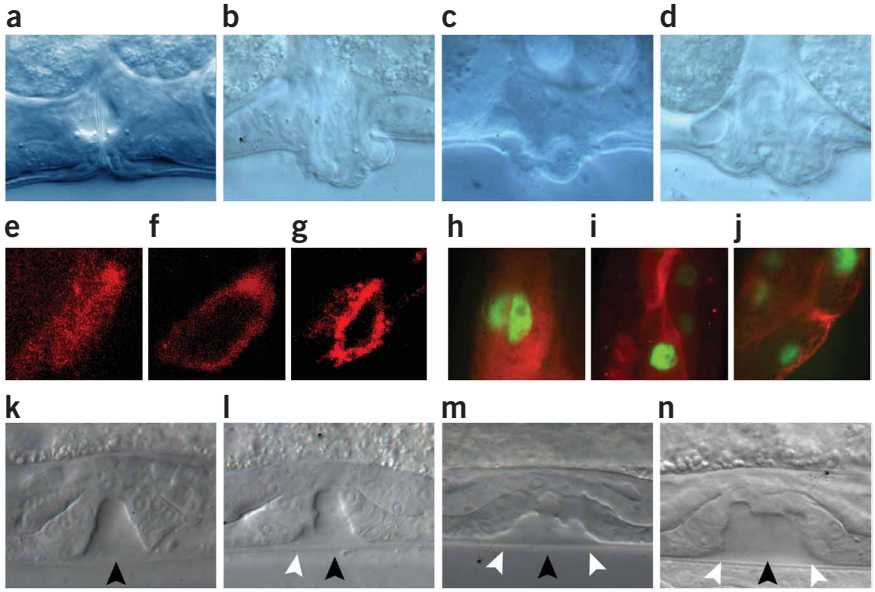

Figure 4 Consequences of SHOC2^{S2G} expression in *C. elegans* vulva development. (**a**–**d**) Nomarski images of vulvas of adult animals. A normal vulva is observed in animals expressing SHOC2^{wt} (**a**), whereas in worms expressing SHOC2^{S2G} (**b**,**c**) or myr::SHOC2^{wt} (**d**) a protrusion of the vulva is visible. (**e**–**j**) Subcellular localization of V5-tagged SHOC2 proteins in *C. elegans* cells. In excretory canal cells (**e**–**g**) and intestinal cells (**h**–**j**), SHOC2^{wt} protein is present throughout the cytoplasm (**e**,**h**), whereas both SHOC2^{S2G} (**f**,**i**) and myr::SHOC2^{wt} (**g**,**j**) are enriched in or restricted to the plasma membrane. Anti-V5 (red) was used to visualize SHOC2 proteins. In intestinal cells, nuclei express GFP due to pelt-2::GFP plasmid used as a marker for transformation. (**k**–**n**) Nomarski images of vulval precursor cells at the L3 stage. In animals expressing SHOC2^{wt} only P6.p descendants invaginate (**k**), whereas in SHOC2^{S2G}- (**l**,**m**) and myr::SHOC2^{wt}-expressing (**n**) animals, P5.p (**l**–**n**) and P7.p descendants (**m**,**n**) also detach from the cuticle. Black arrowheads point to P6.p descendant invagination, whereas white arrowheads point to P5.p and P7.p descendant invagination. Anterior is to the left and dorsal is up in all images.

the disease-causing SHOC2^{S2G} is expected to promote prolonged PP1C-mediated RAF1 dephosphorylation at Ser259 and, consequently, sustained RAF1-stimulated MAPK activation, which is consistent with our findings.

In *C. elegans*, N-myristoylated SHOC2 expression altered morphogenesis during vulval development, a process for which the involvement of Ras signaling is well established. Specification of VPCs, however, was not altered. Rather, perturbation of the morphogenetic movements of the VPC descendant cells was observed. Although numerous mutations that alter vulval specification and morphogenesis have been identified, far less is known about processes affecting only morphogenesis[15,16]. It is possible that SHOC2^{S2G} alters RAS signaling in steps downstream of the induction of the vulval fate. Alternatively, SHOC2^{S2G}-induced vulva defects might arise through perturbation of signaling pathways other than RAS-MAPK, such as signaling mediated by the Rho GTPase, Rac, which are critical for vulval morphogenesis[17].

A unique feature of the *SHOC2* mutation is its association with loose anagen hair (LAH). This phenotype occurs in isolation or with Noonan syndrome and has been without molecular cause. Hair shafts from affected individuals show features of the anagen stage of hair follicle development, during which epithelial stem cells proliferate in the hair bulb[18]. In individuals with LAH however, hair bulbs lack internal and external root sheaths. Our findings suggest perturbation in the proliferation, survival or differentiation of epithelial stem cell–derived cells residing in hair follicles and implicate SHOC2-mediated signal transduction in this aspect of stem cell biology.

Finally, we successfully used the human interactome and a network-based statistical method to predict the involvement of a gene in human disease. Our leading candidate, *SHOC2*, was a relatively obscure gene that caused no distinctive phenotype when mutated in worms, providing evidence of the strength of this approach. In future projects, it is anticipated that successful candidates will not be deemed this favorable, necessitating resequencing of many low-probability candidate genes. Emerging interactome datasets and improved analytic methods are likely to enhance the predictive power of systems biology.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Accession Codes. GenBank: *SHOC2*, NM_007373.3; SHOC2, NP_031399.2. Ensembl: *SHOC2*, ENSG00000108061.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are indebted to the affected individuals and families who participated in the study, the physicians who referred the subjects, A. Fire (Stanford University School of Medicine, Stanford, California) and J.D. McGhee (University of Calgary, Calgary, Canada) for plasmids, and C. Ramoni, S. Venanzi and T. Squatriti (Istituto Superiore di Sanità, Rome, Italy) and the Open Laboratory (IGB-CNR, Naples, Italy) for experimental support. Some nematode strains used in this work were provided by the *Caenorhabditis* Genetics Center (University of Minnesota, Minneapolis, Minnesota), funded by the US National Institutes of Health (NIH) National Center for Research Resources. We also thank M.C. Silengo (Università di Torino, Turin, Italy), S. Spranger (Praxis fuer Humangenetik, Bremen, Germany), I.M. Gaspar (Egas Moniz Hospital, Lisbon, Portugal) and D.R. Bertola (HC/FMUSP, São Paulo, Brazil) for their contribution in DNA sampling and valuable clinical assistance. This research was funded by grants from Telethon-Italy (GGP07115) and 'Convenzione Italia-USA-malattie rare' to M.T., the NIH (HL71207, HD01294 and HL074728) to B.D.G., SBCNY (P50GM071558) to A.M. and R.I. the German Research Foundation (DFG) (ZE 524/4-1) to M.Z. and IRCCS-CSS (Ricerca Corrente 2009) to F.L.

AUTHOR CONTRIBUTIONS

V.C. and V.F. were responsible for mutation analysis and biochemistry, E.D.S., S.M. and P.B. were responsible for the generation and phenotypic characterization of transgenic *C. elegans* strains, L.A.P., J.M., W.S. and A.L. performed high-throughput resequencing, A.S., F.L. and C.R. did mutation analysis, A.M. and R.I. did protein network analysis, S.C. did the confocal laser scanning microscopy, A.C., E.F. and D.M. were responsible for functional studies, G.Z., L.M., M.C.D., D.B., K.K., C.A., A.S., G.B.F., R.T., M.Z. and B.D. obtained DNA specimens from patients and did clinical evaluation, and B.D.G. and M.T. were responsible for the project planning, data analysis and preparation of manuscript.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Resh, M.D. Trafficking and signaling by fatty-acylated and prenylated proteins. *Nat. Chem. Biol.* **2**, 584–590 (2006).
2. Farazi, T.A., Waksman, G. & Gordon, J.I. The biology and enzymology of protein N-myristoylation. *J. Biol. Chem.* **276**, 39501–39504 (2001).
3. Selfors, L.M., Schutzman, J.L., Borland, C.Z. & Stern, M.J. soc-2 encodes a leucine-rich repeat protein implicated in fibroblast growth factor receptor signaling. *Proc. Natl. Acad. Sci. USA* **95**, 6903–6908 (1998).
4. Sieburth, D.S., Sun, Q. & Han, M. SUR-8, a conserved Ras-binding protein with leucine-rich repeats, positively regulates Ras-mediated signaling in *C. elegans*. *Cell* **94**, 119–130 (1998).
5. Li, W., Han, M. & Guan, K.L. The leucine-rich repeat protein SUR-8 enhances MAP kinase activation and forms a complex with Ras and Raf. *Genes Dev.* **14**, 895–900 (2000).
6. Rodriguez-Viciana, P., Oses-Prieto, J., Burlingame, A., Fried, M. & McCormick, F. A phosphatase holoenzyme comprised of Shoc2/Sur8 and the catalytic subunit of PP1 functions as an M-Ras effector to modulate Raf activity. *Mol. Cell* **22**, 217–230 (2006).
7. Mazzanti, L. *et al.* Noonan-like syndrome with loose anagen hair: a new syndrome? *Am. J. Med. Genet. A.* **118A**, 279–286 (2003).
8. Schubbert, S., Shannon, K. & Bollag, G. Hyperactive Ras in developmental disorders and cancer. *Nat. Rev. Cancer* **7**, 295–308 (2007).
9. Tartaglia, M. & Gelb, B.D. Molecular genetics of NS. in *Monographs in Human Genetics, Vol. 17. NS and Related Disorders: A Matter of Deregulated RAS Signaling* (ed. Zenker, M.) 20–39 (Karger Press, Basel, Switzerland, 2009).
10. Berger, S.I., Posner, J.M. & Ma'ayan, A. Genes2Networks: connecting lists of gene symbols using mammalian protein interactions databases. *BMC Bioinformatics* **8**, 372 (2007).
11. Boutin, J.A. Myristoylation. *Cell. Signal.* **9**, 15–35 (1997).
12. Tartaglia, M. *et al.* Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat. Genet.* **34**, 148–150 (2003).
13. Araki, T. *et al.* Mouse model of NS reveals cell type- and gene dosage-dependent effects of Ptpn11 mutation. *Nat. Med.* **10**, 849–857 (2004).
14. Loh, M.L. *et al.* Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. *Blood* **103**, 2325–2331 (2004).
15. Sternberg, P.W. Vulval development. in *Wormbook* (ed. The *C. elegans* Research Community) doi/10.1895/wormbook.1.6.1 http://www.wormbook.org (2005).
16. Eisenmann, D.M. & Kim, S.K. Protruding vulva mutants identify novel loci and Wnt signaling factors that function during *Caenorhabditis elegans* vulva development. *Genetics* **156**, 1097–1116 (2000).
17. Kishore, R.S. & Sundaram, M.V. ced-10 Rac and mig-2 function redundantly and act with unc-73 trio to control the orientation of vulval cell divisions and migrations in *Caenorhabditis elegans*. *Dev. Biol.* **241**, 339–348 (2002).
18. Tosti, A. *et al.* Loose anagen hair in a child with Noonan's syndrome. *Dermatologica* **182**, 247–249 (1991).

ONLINE METHODS

Constructing a mammalian protein-protein interaction network from available resources. The protein-protein and signaling networks we chose were all literature-based 'legacy' direct biochemical mammalian interactions from low-throughput functional experiments that had been extracted manually (literature curated). We did not include interactions from high-throughput methods, orthologous interactions from lower organisms or interactions predicted using *in silico* methods. We considered only direct biophysical binding or enzymatic interactions and excluded interactions based on functional association. The following available protein-protein interaction datasets were used: DIP[19] (updated to 30 May 2006); IntAct[20] (updated to 12 June 2006); MINT[21] (updated to 21 May 2006); Ma'ayan *et al.*[22] (updated to 21 May 2006); BIND[23] (updated to 24 January 2006); PDZBase[24] (updated to 25 September 2006). We choose these datasets because components in those networks were annotated with accession codes that permit data consolidation and those datasets were provided freely for analysis and reuse. All interactions from these databases are claimed to be direct biochemical interactions determined experimentally, and all include the PubMed reference for the research article that described the experiments used to identify the interactions. Consolidating interactions from the different network databases was accomplished by combining human, mouse and rat gene symbols using the xml version of Swiss Prot (21 June 2006).

Algorithm used to generate a list of Noonan syndrome candidate genes. We began by considering the following problem: given a graph G in which a small subset of vertices S, $S \subset G$, are identified as seed nodes, in this case known disease genes that cause Noonan syndrome, find a close to minimum connected subgraph G' that includes the seed nodes in S while pruning out intermediate nodes and links that are not statistically significant for interacting with the seed list. In order to consider this problem, we used an algorithm with six instructional steps. (i) Combine available mammalian protein-protein interaction networks using as described[10]. (ii) Filter the merged network to prune out interactions from publications reporting high-throughput interaction data as described[10]. (iii) Find all shortest paths[25] of length k_1 between all pairs of vertices in the merged seed list $S' \cup S''$ of all known Noonan syndrome disease genes. (iv) Find all edges between intermediate vertices identified in (iii). Intermediate vertices, I, are vertices that fall on shortest paths between pairs between all pairs in $S' \cup S''$ such that $I \subset G$ and $I \notin \{S' \cup S''\}$. (v) Combine all nodes and links found in (iii) and (iv) to create the subnetwork G'. (vi) Rank intermediates base on their links in background network versus links in subnetwork using a binomial proportions test as described[10].

Subjects and mutation analysis. Genomic DNAs from a cohort of 96 subjects with Noonan syndrome or a phenotype suggestive of this disorder without mutation in previously identified disease genes (*PTPN11*, *SOS1*, *KRAS*, *HRAS*, *RAF1*, *BRAF*, *MEK1* and *MEK2*) were screened for the entire *SHOC2* coding region using high-throughput resequencing as previously described[26]. All sequence variants identified were verified by manual inspection of the chromatograms and putative causative mutations were verified using another independent sequencing reaction. *SHOC2* was then analyzed in a panel of 410 mutation-negative individuals with Noonan syndrome or a clinically related phenotype by denaturing high-performance liquid chromatography and direct sequencing[26]. In this cohort, clinical features for the majority of subjects satisfied standardized diagnostic criteria[27–31], but a few individuals who lacked sufficient features for a definitive diagnosis were also included. DNA from skin fibroblasts, hair bulbs and/or epithelial cells from the oral mucosa was extracted using standard protocols. Samples were collected under research projects approved by the Institutional Review Boards at the Istituto Superiore di Sanità, Mount Sinai School of Medicine, IRCCS-Casa Sollievo della Sofferenza, Università di Bologna, Ospedale 'Bambino Gesù', Università di Torino, University of Erlangen-Nuremberg and Universitatsklinikum Hamburg-Eppendorf, with informed consent. Permission was obtained to publish the photographs of subjects shown in **Figure 1**. When available, parental DNAs were sequenced to establish whether identified changes were *de novo*. Paternity was confirmed using the AmpF/STR Identifier PCR Amplification Kit (Applied Biosystems).

Functional analyses. *In silico* analysis of protein N-myristoylation was performed using the Myristoylator, TermiNator and NMT softwares available online. The nucleotide substitutions of interest were introduced in V5- and Myc-tagged (C terminus) human *SHOC2* cDNA expression constructs by site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Stratagene). Cos-1, 293-T and Neuro2A cells were maintained in DMEM (Gibco) supplemented with 10% heat-inactivated FBS (Euroclone) and antibiotics and were transfected at 60–70% confluency using Fugene6 (Roche) or Lipofectamine 2000 (Invitrogen). N-myristoylation was evaluated by [^{3}H]myristic acid (30 μCi/ml) incorporation as described elsewhere[32]. Proteins immunoprecipitated with an anti-V5 antibody from cell lysates were separated by SDS-PAGE. Gels were fixed, soaked in Amplify (Perkin Elmer) for 30 min, dried under a GelAir drying frame (Bio-Rad) and exposed to X-ray film (Kodak) for 2 months. Cellular fractionation and ERK phosphorylation assays were performed on Cos-1 cells transiently expressing the V5-tagged SHOC2^{wt} or SHOC2^{S2G} using standard protocols[26,33]. Cells were serum starved for 16 h and then stimulated with EGF (30–100 ng/ml) for the indicated intervals. In all experiments, a human *NMT1* cDNA expression construct (Origene) was co-transfected to ensure that the amount of endogenous NMT would not be limiting.

Confocal laser scanning microscopy. 3×10^3 cells were seeded on glass coverslips, transiently transfected, serum starved for 16 h and stimulated with EGF (30 ng/ml, 15 min). Cells were fixed with 3% paraformaldehyde (30 min, 4 °C), permeabilized with 0.5% Triton X-100 (10 min, 25 °C), and stained as described in the figure legends. Imaging was performed on a Leica TCS SP2 AOBS apparatus, using excitation spectral laser lines at 405, 488 and 594 nm, tuned with an acousto-optical tunable filter. Image acquisition and processing were conducted by using the Leica Confocal Software (Leica Lasertechnik GmbH). Signals from different fluorescent probes were taken in sequential scanning mode.

Generation of *C. elegans* strains and phenotypic analysis. Culture, maintenance and genetic crosses for nematodes were as described[34]. Nematode strains were provided by the *Caenorhabditis* Genetics Center (University of Minnesota). The following mutant alleles were used: *sur-8^{rf}: sur-8(ku167) IV; let-60^{gof}: let-60(n1046) IV; let-23^{rf}: let-23(sy1) II.* V5-tagged *SHOC2^{wt}* and *SHOC2^{S2G}* cDNA were subcloned into the heat shock–inducible pPD49.83 vector (a gift of A. Fire, Stanford University School of Medicine). A chimeric SHOC2 protein, myr::SHOC2^{wt}, in which the first seven amino acid residues were substituted by the N-terminal myristoylation signal (MGSCIGK) of *src-2* was obtained via PCR amplification and cloned into the pPD49.83 vector. Germline transformation was performed as described[35]. *elt-2*::GFP (pJM67, a gift from J.D. McGhee, University of Calgary), which drives GFP expression in intestinal cells, was used as co-injection marker. At least three independent lines for each construct were tested for the Pvl phenotype after heat shock. All the lines expressing SHOC2^{S2G} or myr::SHOC2^{wt} upon heat shock showed a Pvl phenotype. Only the lines carrying the following transgenes were scored quantitatively at the compound microscope and used for further analyses and crosses: *gbEx240*[*hsp16.2*::*SHOC2^{wt}*::*V5*; *pelt-2*::*GFP*], *gbEx208a*[*hsp16.2*::*SHOC2^{S2G}*::*V5*; *pelt-2*::*GFP*] and *gbEx209*[*hsp16.2*::*myr*::*SHOC2^{wt}*::*V5*; *pelt-2*::*GFP*]. Genetic crosses were performed according to standard methods. The presence of *sur-8(ku167), let-60(n1046)* and *let-23(sy1)* alleles was confirmed by sequencing the appropriate region of genomic DNA from each transgenic strain. After each cross, isogenic worms that had lost the transgene were cloned separately and used as controls. Animals were scored blindly at the dissecting microscope to count the number of eggs *in utero* after cutting the mother (Egl), to identify animals that had become bags of worms (Bag) and to check for the presence of multiple ectopic pseudovulvae (Muv). A subset of worms was also scored blindly at the compound microscope for vulva morphology and VPC induction phenotypes.

***C. elegans* heat shock experiments, microscopy and immunocytochemistry.** At different developmental stages, worms carrying the transgenes were subjected to heat shock at 33 °C for 30 min and then kept at 30 °C for 1 h. Synchronized embryos were heat shocked to study the effects of transgene expression on embryonic and early larval development, while synchronized L1 and L2 larvae were heat shocked to study the effects on later larval development, movement and fertility. To study VPC induction and vulva morphogenesis, hermaphrodites were heat shocked at early L3 stages and animals were scored for vulval induction at the L4 stage and for Pvl phenotype

doi:10.1038/ng.425

at the adult stage. Microscopy observations were performed with a Zeiss Axioskop equipped with epifluorescence and differential interference contrast on live animals anesthetized and mounted on 2% agarose pads containing 10 mM sodium azide. Images were collected with an Axiocam digital camera. Confocal analyses were performed using a Leica TCS SP2 microscope. For immunocytochemistry analyses, transgenic worms were heat shocked for 2 h and then were fixed with 2% PFA (25 °C, 5 min, 1 h on ice). They were processed as reported[36] and then incubated overnight in a dilution of anti-V5 monoclonal (1:200). After repeated washing for 24 h, animals were incubated overnight with secondary Texas Red–conjugated anti-mouse (1:100) (Invitrogen), washed and mounted for observation on microscope slides.

URLs. DIP, http://dip.doe-mbi.ucla.edu/; IntAct, ftp://ftp.ebi.ac.uk/pub/databases/intact/current; MINT, http://mint.bio.uniroma2.it/mint/; Iyengar web resources, http://www.mssm.edu/labs/iyengar/resources; BIND, http://www.bind.ca/; PDZBase, http://icb.med.cornell.edu/services/pdz/start; Prosite, http://www.expasy.ch/tools/scanprosite/; Pfam database, http://pfam.janelia.org/; Myristoylator, http://www.expasy.org/tools/myristoylator/; TermiNator, http://www.isv.cnrs-gif.fr/terminator3/index.html; NMT, http://mendel.imp.ac.at/myristate/SUPLpredictor.htm.

19. Xenarios, I. *et al.* The Database of Interacting Proteins: 2001 update. *Nucleic Acids Res.* **29**, 239–241 (2001).
20. Kerrien, S. *et al.* IntAct-open source resource for molecular interaction data. *Nucleic Acids Res.* **35**, D561–D565 (2007).
21. Chatr-aryamontri, A. *et al.* MINT: the Molecular INTeraction database. *Nucleic Acids Res.* **35**, D572–D574 (2007).
22. Ma'ayan, A. *et al.* Formation of regulatory patterns during signal propagation in a Mammalian cellular network. *Science* **309**, 1078–1083 (2005).
23. Bader, G.D., Betel, D. & Hogue, C.W.V. BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* **31**, 248–250 (2003).
24. Beuming, T., Skrabanek, L., Niv, M.Y., Mukherjee, P. & Weinstein, H. PDZBase: a protein-protein interaction database for PDZ-domains. *Bioinformatics* **21**, 827–828 (2005).
25. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271 (1959).
26. Tartaglia, M. *et al.* Gain-of-function SOS1 mutations cause a distinctive form of NS. *Nat. Genet.* **39**, 75–79 (2007).
27. Allanson, J.E. Noonan syndrome. *J. Med. Genet.* **24**, 9–13 (1987).
28. van der Burgt, I. *et al.* Clinical and molecular studies in a large Dutch family with Noonan syndrome. *Am. J. Med. Genet.* **53**, 187–191 (1994).
29. Voron, D.A., Hatfield, H.H. & Kalkhoff, R.K. Multiple lentigines syndrome. Case report and review of the literature. *Am. J. Med.* **60**, 447–456 (1976).
30. Sarkozy, A. *et al.* LEOPARD syndrome: clinical aspects and molecular pathogenesis. in *Monographs in Human Genetics, Vol 17. NS and Related Disorders: A Matter of Deregulated RAS Signaling* (ed. Zenker, M.) 55–65 (Karger Press, Basel, Switzerland, 2009).
31. Roberts, A. *et al.* The cardiofaciocutaneous syndrome. *J. Med. Genet.* **43**, 833–842 (2006).
32. Utsumi, T. *et al.* Amino acid residue penultimate to the amino-terminal Gly residue strongly affects two cotranslational protein modifications, *N*-myristoylation and *N*-acetylation. *J. Biol. Chem.* **276**, 10505–10513 (2001).
33. Vachon, L., Costa, T. & Hertz, A. GTPase and adenylate cyclase desensitize at different rates in NG108–15 cells. *Mol. Pharmacol.* **31**, 159–168 (1987).
34. Sulston, J.E. & Hodgkin, J. Methods. in *The Nematode Caenorhabditis elegans* (ed. Wood, W.B. and The Community of *C. elegans* Researchers) 587–606 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 1988).
35. Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans* after microinjection of DNA into germline cytoplasm: extrachromosomal maintenance and intergration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
36. Duerr, J.S. Immunohistochemistry. in *WormBook* (ed. The *C. elegans* Research Community) doi/10.1895/wormbook.1.105.1 (2006).